

REMARKS

This Amendment is in response to the Office Action mailed October 3, 2002. In the Office Action, claims 2-6, 8-17 and 20-24 have been rejected under 35 U.S.C. §103(a). Applicant respectfully disagrees with the rejections in their entirety. However, for this application, claims 6 and 12 have been revised to include the limitations of dependent claims 2-3 and claim 8, respectively. Therefore, claims 6 and 12 are substantially equivalent to prior pending claims 3 and 8. Claim 15 and 22 have been placed into independent format to include limitations of base claims 13-14 and 20-21, respectively. Claims 25-37 have been added. No new search is necessary since the revised independent claims are in condition for allowance.

I. REJECTION OF CLAIMS 6, 12, 13 AND 20

Claims 6, 12, 13 and 20 have been rejected under 35 U.S.C. §103(a) as being unpatentable over U.S. Patent No. 5,311,588 issued to Polcyn, et al. (Polcyn) in view of U.S. Patent No. 6,125,645 issued to Modi, et al. (Modi). Applicant traverses the rejection in its entirety and incorporates the arguments set forth in the Response dated June 28, 2002. However, further discussion of these claims is not warranted in light of the revisions set forth to claim 6 and 12 and the cancellation, without prejudice, of claims 13 and 20.

II. REJECTION OF CLAIMS 2-4, 8-9, 11, 14-17 AND 21-24

Claims 2-4, 8-9, 11, 14-17 and 21-24 have been rejected under 35 U.S.C. §103(a) as being unpatentable over Polcyn and Modi in view of U.S. Patent No. 6,223,154 issued to Nicholls et al. (Nicholls). The Office Action states that both Polcyn and Modi do not teach using short-term averaged energy, long-term averaged energy, etc. in the voice/speech detection process. *See Paragraph 6 of the Office Action.* However, the Office Action alleges that Nicholls discloses a voice activity detector "which calculates average energy values of

received frames and determines if the average value exceeds a threshold and compares a staged average energy value to a current frame energy value." *See Paragraph 6 of the Office Action.* Applicant respectfully disagrees.

The teachings of Nicholls are directed to the usage of staggered average value for the clipping of speech during speakerphone communication. Such teachings fail to describe or even suggestion the use of short-term and long-term averaged energies for determining silence or voice signaling as set forth in claims 6, 12, 15 and 22. Moreover, Nicholls fails to describe or suggest a tiered evaluation of reviewing short-term and long-term averaged energies prior to a determination of the peak-to-mean likelihood ratio (PMLR) as now set forth in claims 6, 12 and 22. The PMLR is calculated only if the sum of the short-term averaged energy and a factor is greater than the long-term averaged energy (current audio frame is not silence).

In light of the foregoing, Applicant respectfully requests that the outstanding §103(a) rejection be withdrawn.

VERSION WITH MARKINGS TO SHOW CHANGES MADE

This complete set of claims excludes claims 2-3, 8, 13-14 and 20-21 and includes revised claims 4-6, 9, 12, 15, 22 and 24. New dependent claims 25-37 have been added.

1 1. Cancelled.

1 2. Cancelled.

1 3. Cancelled.

1 4. (Amended) The method of claim 6[3], upon determining that the sum is
2 greater than the long-term averaged energy and before determining the peak-to-mean
3 likelihood ratio, the method further comprises:
4 determining whether a difference between the long-term averaged energy and the
5 short-term averaged energy is less than a predetermined threshold;
6 determining that the current audio frame represents voice if the difference is greater
7 than the predetermined threshold; and
8 continuing by determining the peak-to-mean likelihood ratio if the difference is less
9 than the predetermined threshold.

1 5. (Amended) The method of claim 6[2], wherein the determining of the short-
2 term averaged energy comprises:
3 determining an energy, in decibels, of the current audio frame;
4 determining a short-term averaged energy for a prior audio frame; and
5 conducting a weighted average of the energy of the current audio frame and the short-
6 term averaged energy for the prior audio frame.

1 6. (Three Times Amended) A method for enhancing voice activity detection
2 comprising:

3 determining a short-term averaged energy for a current audio frame;
4 determining a long-term averaged energy for the current audio frame;
5 determining whether a sum of the short-term averaged energy and a factor is greater
6 than the long-term averaged energy;
7 determining that the current audio frame represents silence if the sum is less than the
8 long-term averaged energy, without necessitating a determination of the peak-to-mean
9 likelihood ratio;
10 determining a peak-to-mean likelihood ratio, the determining a peak-to-mean
11 likelihood ratio comprises
12 calculating an averaged peak-to-mean ratio for the current audio frame,
13 determining a maximum averaged peak-to-mean ratio,
14 determining a minimum averaged peak-to-mean ratio,
15 determining a difference between the maximum averaged peak-to-mean ratio
16 and the averaged peak-to-mean ratio for the current audio frame,
17 determining a difference between the maximum averaged peak-to-mean ratio
18 and the minimum averaged peak-to-mean ratio, and
19 conducting a ratio, a denominator of the ratio being the difference between the
20 maximum averaged peak-to-mean ratio and the minimum averaged peak-to-mean
21 ratio, the numerator being the difference between the maximum averaged peak-to-
22 mean ratio and the averaged peak-to-mean ratio; and
23 comparing the peak-to-mean likelihood ratio to a selected threshold to determine
24 whether the[a] current audio frame represents a voice signal.

1 7. Cancelled.

1 8. Cancelled.

1 9. (Amended) The communication module of claim 12[8], wherein the voice
2 activity detector, when executed, controls the processing unit to determine whether a
3 difference between the long-term averaged energy and the short-term averaged energy is less
4 than a predetermined threshold, and to signal that the current audio frame represents voice if
5 the difference is greater than the predetermined threshold.

1 10. Cancelled.

1 11. (Amended) The communication module of claim 9, wherein the voice activity
2 detector, when executed, controls the processing unit to determine a peak-to-mean ratio by (i)
3 sampling an analog signal a predetermined number of times to produce a plurality of sampled
4 signals each having a sampled value, (ii) determining a maximum value of the plurality of
5 sampled signals, and (iii) conducting a ratio between an absolute value of the maximum
6 value and a summation of the sampled values for the plurality of sampled signals.

1 12. (Twice Amended) A communication module comprising:
2 a substrate;
3 a processing unit placed on the substrate; and
4 a memory coupled to the processing unit, the memory to contain a voice activity
5 detector which, when executed, controls the processing unit to
6 determine whether a sum of a short-term averaged energy and a
7 predetermined factor is greater than a long-term averaged energy, and to signal that a
8 current audio frame represents silence if the sum is less than the long-term averaged
9 energy, and
10 if the current audio frame is not determined to be silence using the short-term
11 averaged energy and the long-term averaged energy, determine a peak-to-mean

12 likelihood ratio for the current audio frame by (i) monitoring a maximum averaged
13 peak-to-mean ratio and a minimum averaged peak-to-mean ratio, (ii) determining a
14 first result being a difference between the maximum averaged peak-to-mean ratio and
15 the averaged peak-to-mean ratio for the current audio frame, (iii) determining a
16 second result being a difference between the maximum averaged peak-to-mean ratio
17 and the minimum averaged peak-to-mean ratio, and (iv) conducting a ratio between
18 the first result as a numerator and the second result as a denominator₁[;] and
19 comparing₁[e] the peak-to-mean likelihood ratio₁[n] to a selected threshold to
20 determine whether the current audio frame represents a voice signal.

1 13. Cancelled.

1 14. Cancelled.

1 15. (Amended) A[The] machine readable medium having embodied thereon a
2 computer program for processing by a machine. [of claim 14, wherein] the computer
3 program [further] comprising:
4 a first routine for determining a normalized peak-to-mean likelihood ratio including
5 (i) a denominator having a value substantially equal to a difference between a maximum
6 averaged peak-to-mean ratio and a minimum averaged peak-to-mean ratio and (ii) a
7 numerator having a value substantially equal to a difference between the maximum averaged
8 peak-to-mean ratio and the averaged peak-to-mean ratio;
9 a second routine for comparing the peak-to-mean likelihood ratio to a selected
10 threshold to determine whether a current audio frame being transmitted represents a voice
11 signal;

12 a third routine for determining a short-term averaged energy for successive audio
13 frames including the current audio frame, the third routine being executed before the first and
14 second routines;
15 a fourth routine for determining a long-term averaged energy for the current audio
16 frame, the fourth routine being executed before the first and second routines;
17 a fifth routine for determining whether a sum of the short-term averaged energy and a
18 predetermined factor is greater than the long-term averaged energy, the fifth routine being
19 executed before the first and second routines; and
20 a sixth routine for determining whether a difference between the long-term averaged
21 energy and the short-term averaged energy is less than a predetermined threshold, the sixth
22 routine being executed after determining that the sum is greater than the long-term averaged
23 energy and before execution of the first and second routines.

1 16. The machine readable medium of claim 15, wherein the fifth routine
2 determining that the current audio frame represents silence if the sum is less than the long-
3 term averaged energy.

1 17. The machine readable medium of claim 15, wherein the sixth routine
2 determining that the current audio frame represents voice if the difference is greater than the
3 predetermined threshold.

1 18. Cancelled.

1 19. Cancelled.

1 20. Cancelled.

1 21. Cancelled.

1 22. (Amended) A [The] method for enhancing voice activity detection [of claim
2 21, wherein after determining the short-term averaged energy and the long-term averaged

3 energy, the method further comprises] comprising:

4 determining a short-term averaged energy for a current audio frame;

5 determining a long-term averaged energy for the current audio frame;
6 determining whether a sum of the short-term averaged energy and a factor is greater
7 than the long-term averaged energy; [and]

8 determining that the current audio frame represents silence if the sum is less than the
9 long-term averaged energy, without necessitating a determination of the peak-to-mean
10 likelihood ratio;

11 determining a peak-to-mean likelihood ratio including (i) a denominator having a
12 value substantially equal to a difference between a maximum averaged peak-to-mean ratio
13 and a minimum averaged peak-to-mean ratio and (ii) a numerator having a value
14 substantially equal to a difference between the maximum averaged peak-to-mean ratio and
15 the averaged peak-to-mean ratio; and

16 comparing the peak-to-mean likelihood ratio to a selected threshold to determine
17 whether a current audio frame represents a voice signal.

1 23. The method of claim 22, upon determining that the sum is greater than the
2 long-term averaged energy and before determining the peak-to-mean likelihood ratio, the
3 method further comprises:

4 determining whether a difference between the long-term averaged energy and the
5 short-term averaged energy is less than a predetermined threshold;

6 determining that the current audio frame represents voice if the difference is greater
7 than the predetermined threshold; and

8 continuing by determining the peak-to-mean likelihood ratio if the difference is less
9 than the predetermined threshold.

1 24. (Amended) The method of claim 22[21], wherein the determining of the
2 short-term averaged energy comprises:
3 determining an energy, in decibels, of the current audio frame;
4 determining a short-term averaged energy for a prior audio frame; and
5 conducting a weighted average of the energy of the current audio frame and the short-
6 term averaged energy for the prior audio frame.

1 25. (New) The method of claim 6, wherein the short-term averaged energy is an
2 accumulation of signal energy associated with successive audio frames including the current
3 audio frame.

1 26. (New) The method of claim 25, wherein the successive audio frames are
2 pulse code modulation (PCM) audio frames.

1 27. (New) The method of claim 25, wherein the long-term averaged energy is
2 based on the accumulation of the signal energy and a background noise level.

1 28. (New) The method of claim 6, wherein the short-term averaged energy is
2 based on a current frame entry and a prior short-term averaged energy value.

1 29. (New) The method of claim 6, wherein the factor is at least two decibels.

1 30. (New) The communication module of claim 12, wherein the short-term
2 averaged energy determined by the voice activity detector is an accumulation of signal
3 energy associated with the successive audio frames being pulse code modulation (PCM)
4 audio frames.

1 31. (New) The communication module of claim 30, wherein the long-term
2 averaged energy determined by the voice activity detector is based on the accumulation of
3 the signal energy and a background noise level.

1 32. (New) The communication module of claim 12, wherein the predetermined
2 factor is at least two decibels.

1 33. (New) The software readable medium of claim 15, wherein the short-term
2 averaged energy determined by the third routine is an accumulation of signal energy
3 associated with the successive audio frames.

1 34. (New) The software readable medium of claim 33, wherein the long-term
2 averaged energy determined by the fourth routine is based on the accumulation of the signal
3 energy and a background noise level.

1 35. (New) The method of claim 22, wherein the short-term averaged energy is an
2 accumulation of signal energy associated with successive audio frames including the current
3 audio frame.

1 36. (New) The method of claim 22, wherein the short-term averaged energy is
2 based on the current audio frame and a prior short-term averaged energy value.

1 37. (New) The method of claim 22, wherein the factor is at least two decibels.

CONCLUSION

In view of the amendments and remarks made above, it is respectfully submitted that all pending claims are in condition for allowance, and such action is respectfully solicited.

Respectfully submitted,

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP

Dated: December 3, 2002



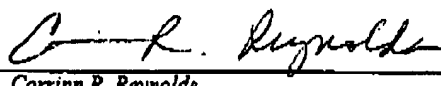
WILLIAM W. SCHAAL

Reg. No. 39,018

12400 Wilshire Boulevard, Seventh Floor
Los Angeles, California 90025
(714) 557-3800

CERTIFICATE OF MAILING

*I hereby certify that this correspondence is being
transmitted via facsimile under 37 CFR §1.8 on:
December 3, 2002.*



Corinn R. Reynolds

12/3/02
Date